AD P001326

# TESTING DURING TRAINING: WHY DOES IT ENHANCE MOVEMENT RETENTION

Joseph D. Hagman
US Army Research Institute for the Behavioral and Social Sciences

Linear movement retention was examined for training methods emphasizing (repeating) either presentation (p) or test (t) trials. P-trials were experimenter-defined study movements constrained by a mechanical stop; t-trials were learner-defined recall movements unconstrained by the stop. Separate groups of governmental employees received training consisting of three, 6-trial cycles. Cycles began with a p-trial that defined the criterion movement to be remembered. The five remaining trials of each cycle varied in type across groups. One group, for example, performed successive t-trials, whereas another performed successive p-trials yoked in value to the first group's t-trials. Retention was then examined at 3 minutes and 24 hours after training. Absolute (unsigned) error revealed that t-trials were more effective than yoked p-trials in promoting movement retention. The data were consistent with the hypothesis that retention benefits obtained from testing during training result from better initial learning (encoding) of kinesthetic cues generated under a learner-defined than under an experimenter-defined movement execution mode. It was concluded that testing cannot only be used to evaluate but also to improve motor skill retention.

TESTING DURING TRAINING:
WHY DOES IT ENHANCE MOVEMENT RETENTION

JOSEPH D. HAGMAN
US ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
ALEXANDRIA, VA 22333

The Army's primary peacetime mission is to maintain combat readiness (Guthrie, 1979). To be combat ready, soldiers must first become proficient in their performance of job tasks, and then, retain this proficiency over what can be prolonged periods of no practice. One way to enable soldiers both to reach and maintain combat readiness is through the use of task training methods that promote effective acquisition and retention. To do this, these methods must be identified and compared.

A review of the training research literature reveals that training methods have been compared primarily within the context of laboratory experiments. Here, training has involved the execution of presentation (p) trials, where to-be-learned information is presented by the experimenter to the learner for study, and test (t) trials, where this information is removed and the learner attempts to recall (reproduce) it from memory. Although standard training methods involve alternation of p- and t-trials (e.g., Tulving, 1967; Wrisberg & Schmidt, 1975), the most effective number and sequential arrangement of p-and t-trials to use is a matter of debate. From a traditional learning theory viewpoint, where p-trials are seen as having an effect similar to reinforcement (Adams & Dijkstra, 1966), training methods that emphasize (repeat) p-trials should be more effective than those that repeat t-trials. P-trial repetition increases the number of reinforcement opportunities during training, and therefore, should enhance both acquisition and retention. From a contemporary cognitive learning viewpoint, on the other hand, information processing activities such as memory retrieval and internal item generation are considered important aspects of acquisition and retention (Bjork, 1975). Because t-trials provide an opportunity to perform these activities on information studied during p-trials, training methods that repeat t-trials should also be effective.

P-trial effects have been documented in numerous experiments showing that improved performance occurs when p-trials are repeated during training (e.g., Adams & Dijkstra, 1966). Only recently, however, have improvements associated with t-trial repetition been reported. Researchers have shown that with verbal tasks t-trials not only contribute to acquisition (e.g., Lachman & Laughery, 1968) but also to retention (Hogan & Kintsch, 1971; Wenger, Thompson, & Bartling, 1980). Even more recently, t-trials have been reported to influence motor task performance. Hagman (1980a,b), for example, had persons learn either the distance (extent) or end-location (terminal position) of linear positioning movements under training methods emphasizing either p- or t-trial repetition. P-trials were movements terminated by a mechanical stop that was prepositioned by the experimenter to ensure execution of the to-be-learned criterion movement cue (i.e., distance or end-location). T-trials were movements performed with the stop removed. It was during t-trials that learners stopped their own movement when they thought they had accurately recalled

the criterion movement cue. Results of both experiments showed that movement cue acquisition was better when p-trials were repeated during training, whereas long-term retention was better when t-trials were repeated during training.

The purpose of the present experiment was to extend these earlier findings by testing two hypotheses suggested (Hagman, 1980) to account for the beneficial effect of t-trials on movement cue retention. The first hypothesis relies on the procedural distinction between experimenter-defined (i.e., performed with the stop present) and learner-defined (i.e., performed with the stop absent) movements. Evidence suggests that movement cues generated under a learner-defined execution mode are retained better than those generated under an experimenter-defined execution mode (Kelso, 1977; Stelmach, Kelso & McCullagh, 1976). This enhanced retention is caused by superior learning (encoding) of learner-defined movement cues brought about by the learner's ability to predict or anticipate cue values prior to movement initiation (e.g., Kelso, 1977). T-trials allow for prediction because they are learner-defined, whereas p-trials do not allow for prediction because they are experimenter-defined. In a multitrial training context learners base posttraining recall attempts on their retention of cues generated during the trial type repeated during training. That is, learners rely on p-trial retention when p-trials are repeated, whereas they rely on t-trial retention when t-trials are repeated. Because t-trials are learner-defined, retention of t-trial generated cues should be superior to retention of p-trial generated cues which are experimenter-defined. Thus, enhanced long-term motor retention should occur with training methods that emphasize learner-defined t-trial repetition.

The second hypothesis proposed to account for the beneficial effect of t-trial repetition on movement cue retention involves the notions of movement variability and motor schema. The motor schema is an abstraction of task and environmental characteristics that develops through repeated and varied movement during training (Schmidt, 1975), and serves as a rule or concept for movement generation. Researchers have found that as variability increases during training the abstracted schema information becomes increasingly resistant to forgetting (Newell & Shapiro, 1976; Posner & Reele, 1970). In the previous experiments by Hagman (1980a,b), variability during training was generated at t-trials because learners were inconsistent in their recall attempts. In contrast, no variability was generated by p-trials because all were identical in terms of distance (Hagman, 1980a) or end-location (Hagman, 1980). As a result, it could be argued that schema strength was greater after repeated t-trial training than after repeated p-trial training. Thus, one would predict better retention under the former than under the latter training method.

The general approach used in the present experiment to test the validity of these two hypotheses involved yoking separate p-trial training method groups to both the t-trial distance and t-trial end-location groups trained earlier. Yoking involved using a mechanical stop to ensure that p-trials of the yoked groups were identical to the t-trials of the other groups in terms of both distance and end-location. Thus,

yoking afforded the means of equating p- and t-trials in terms of variability during training but allowed the distinction to remain between p- and t-trial execution mode (i.e., experimenter- versus learner-defined). If variability per se during training is the key to enhanced retention of movement cues, then one would expect the retention displayed by the two yoked p-trial groups not to differ from that displayed by the two t-trial groups. If, on the other hand, movement execution mode during training is the key to enhanced retention, then one would expect the two t-trial groups to display retention superior to that of the two yoked p-trial groups.

## Method

### Subjects

Sixty governmental employees volunteered to serve as participants in the experiment. All were members of the professional and clerical staff of the Army Research Institute for the Behavioral and Social Sciences.

### Apparatus

Participants were required to make movements from left to right using a metal slide that ran along a linear track consisting of two stainless steel rods 35 inches (88.9 cm) in length. Two Thompson Ball Bushings supported the slide on the rods which were mounted in parallel on a metal frame 4.25 inches (11 cm) apart and 11 inches (27.94 cm) above the frame base. The base rested on a standard table top 31 inches (78.74 cm) from the floor. A second slide was used by the experimenter to stop movement of the first slide along the track. A pointer attached to the experimenter's side of each slide ran along a meter stick to indicate respective slide position. Additional apparatus included a chin rest to stabilize head position, earphones through which tape-recorded procedural commands were delivered, and a blindfold to eliminate visual cues.

### Design

The experiment contained an acquisition and a retention segment as shown in Figure 1. The acquisition segment consisted of 18 training trials divided into three cycles of six trials each. Cycles contained p- and t-trials. P-trials were experimenter-defined movements terminated by the mechanical stop. The stop was prepositioned by the experimenter to ensure that participants executed (studied) the criterion distance end-location at p-trials and duplicated t-trials at yoked p-trials, i.e., py. T-trials were learner-defined recall movements unconstrained by the mechanical stop. Four training method groups were included in the experiment, i.e., DISTANCE PRESENTATION (DP), DISTANCE TEST (DT), END-LOCATION PRESENTATION (LP), and END-LOCATION TEST (LT). Training methods differed in their emphasis on p- and t-trials performed during each cycle. Group DT performed cycles containing an initial to-be-learned criterion p-trial followed by five successive recall t-trials. Group DP performed cycles containing six successive p-trials. The first was the criterion, but the next five were yoked in distance to the corresponding t-trials of Group DT. Yoking was also applied to the two end-location groups in a similar fashion. Because of this yoking

| | ACQUISITION | | | RETENTION | |
|---|---|---|---|---|---|
| GROUPS | CYCLE 1 | CYCLE 2 | CYCLE 3 | 3 MINUTES | 24 HOURS |
| DP | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | T | T |
| DT | P T T T T T | P T T T T T | P T T T T T | T | T |
| LP | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | P $P_Y$ $P_Y$ $P_Y$ $P_Y$ $P_Y$ | T | T |
| LT | P T T T T T | P T T T T T | P T T T T T | T | T |

*$P_Y$=YOKED

Figure 1. Trial sequence for training method groups at acquisition and retention

procedure, Groups DT and LT were trained before Groups DP and LP. Data from the two yoked PRESENTATION groups were collected in the present experiment, whereas data from the two TEST groups were collected earlier (Hagman, 1980a,b). Although trained at different times, subjects in the two yoked groups were drawn from the same population as those in the two TEST groups.

The retention segment of the experiment consisted of a single t-trial performed by each group at both 3 minutes and 24 hours after acquisition, as shown in Figure 1. Separate 2x2 mixed factorial designs were used to examine distance and end-location cue retention. The between-subjects factor was group (DP, DT, or LP, LT) and the within-subjects factor was retention interval (3 minutes, 24 hours). Fifteen participants were assigned to each of the four training method groups with the constraint that each group contain the same proportion of men and women.

Procedure

Participants were instructed to learn and remember either movement distance or end-location depending on their group. Those in groups DP and LP were also told of the yoking procedure. All participants were then shown a written copy of the trial command sequence that they would be hearing and told the meaning of each command. The p-trial command was "Movement" and the t-trial command was "Recall Movement." Each of these commands was preceded by "Ready" and followed by "Rest." At "Ready" the experimenter grasped the participant's hand and placed it on the handle of the slide. Five seconds later, the participant heard either "Movement" or "Recall Movement" depending on the trial type. At "Movement," participants moved the slide across the track until contacting the mechanical stop. At "Recall Movement," those in Groups DT and LT moved the slide across until they felt that they had recalled the criterion distance or end-location, whereas those in Groups DP and LP moved the slide along until contacting a stop. This stop was prepositioned by the experimenter at the distance or end-location recalled by participants in Groups DT and LT at t-trial execution. Five seconds were allowed for movement execution. During this interval, participants received white noise through earphones to eliminate auditory cues resulting from displacement of the slide. "Rest" marked the beginning of a 10-second interval during which participants removed their hand from the slide and placed it on the table in a predetermined resting position. During rest periods the experimenter recorded recall accuracy to the nearest millimeter (when appropriate) and repositioned the stop in preparation for the next trial. After "Rest," participants heard "Ready" and the command sequence for the next trial began. During the retention segment of the experiment, intervals of 3 minutes and 24 hours were inserted between "Rest" and "Ready." In general, participants were instructed not to count during movements and shown the approximate movement speed (i.e., 125 mm/sec) desired by the experimenter. Prior to making the first movement, participants donned their blindfold and earphones, and then were given a 10-second opportunity to move the slide and get a feel for its movement characteristics.

## Results

Algebraic (signed) and absolute (unsigned) error scores were recorded for each t-trial performed during the retention segment of the experiment. No acquisition data were analyzed because yoking prevented any differences in group performances. Each performance measure was analyzed separately.

Retention was examined using a 2x2 mixed factorial Group (DP, DT or LP, LT) by Retention Interval (3 minutes, 24 hours) analysis of variance (ANOVA). Separate ANOVAs were performed on the algebraic and absolute error scores for the two distance groups (DP, DT) and the two end-location groups (LP, LT). No significant ($P < .05$) effects of interest were found for algebraic error, therefore only absolute error scores are reported.

Distance. Mean absolute error scores are shown in Figure 2. The scores for the two distance groups (i.e., DP, DT) are on the left and those for the two end-location groups (i.e., LP, LT) are on the right. The absolute error ANOVA revealed no significant main effects for a significant groups x retention interval interaction, $F(1,28)=6.85$. The rejection region for this and all other analyses was .05. This interaction resulted from an increase in recall error over time for Group DP and an associated decrease in recall error over time for Group DT. Individual comparisons of simple main effects using the least significant difference method (Carmer & Swanson, 1973) revealed that the Group DP error increase was revealed that 3 minutes after training no difference in recall error existed between Groups DP and DT, whereas 24 hours after training Group DP displayed greater recall error than that of Group DT.

End-location. The absolute error ANOVA, for end-location, revealed a significant main effect of group, $F(1,28)=5.85$, demonstrating greater posttraining recall error for Group LP than for Group LT, and a group x retention interval interaction that approached significance, $F(1,28)=3.11$, $.05 < p < .10$. Although nonsignificant by conventional standards, further analysis of simple main effects associated with this interaction was justified by a priori expectations of training method outcome as indicated by the results obtained for distance cue recall. As shown in Figure 2, the marginal interaction resulted from an increase in recall error after training for Group LP while Group LT error remained almost unchanged. Individual comparisons revealed that the Group LP increase was significant, and that Group LT error was statistically stable. Group recall performance did not differ 3 minutes after training while 24 hours after training Group LP error was significantly greater than Group LT error. Conservatively speaking, the absolute error data for both movement distance and end-location cues reveal that training methods that emphasize testing (i.e., DT, LT) prevent posttraining task retention decrements, whereas those that emphasize presentation produce marked posttraining retention decrements. Thus, even the yoking procedure used in the present experiment to increase movement variability during training was unable to prevent forgetting when p-trials were emphasized.

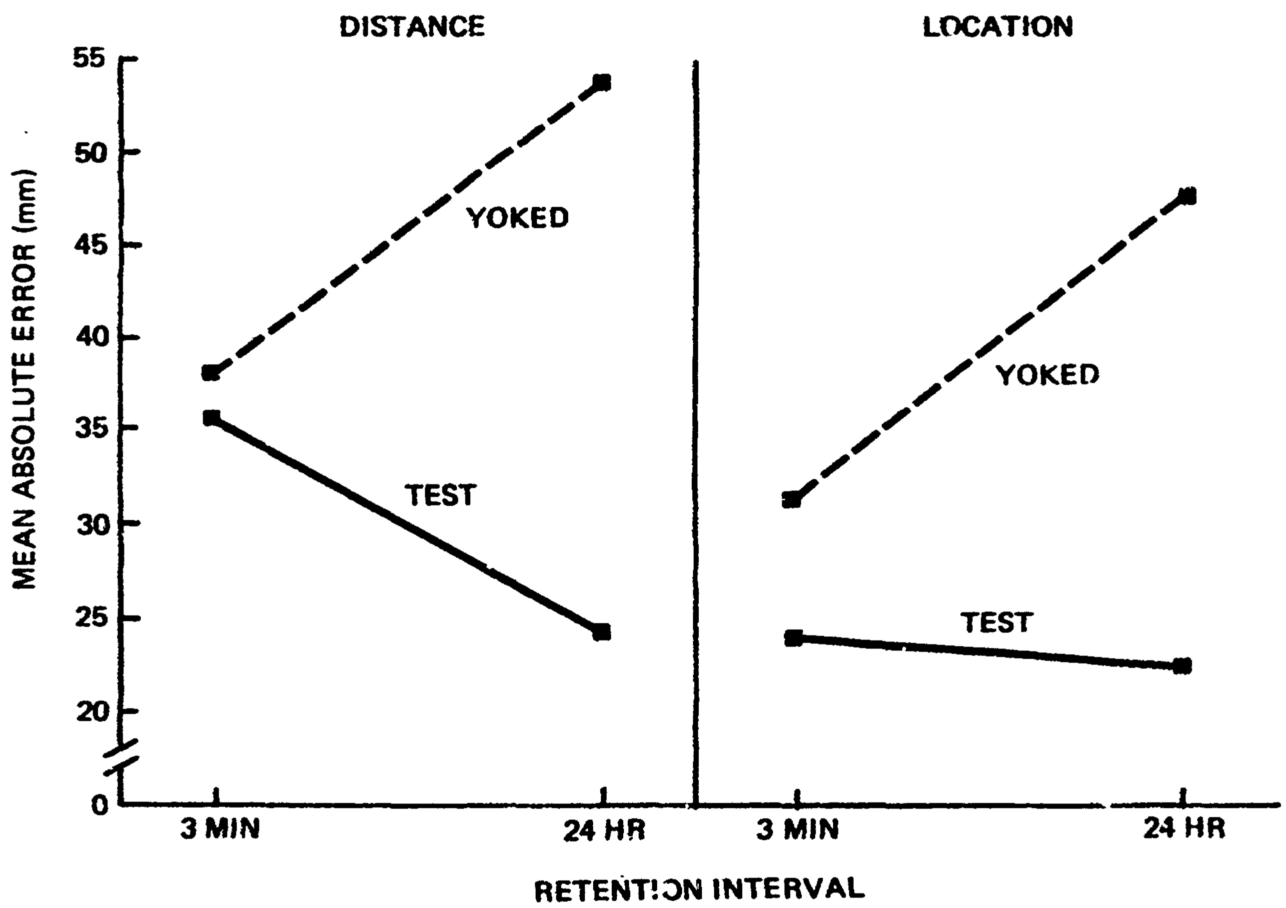


Figure 2. Mean absolute error on retention t-trials for distance and end-location training method groups.

## Discussion

The purpose of this experiment was to explain previous data showing that repeated testing during training is more effective than repeated presentation in promoting long-term motor task retention (Hagman, 1980a,b). Two hypotheses were tested. The first stated that retention benefits were caused by differences in the learning (encoding) characteristics of p- and t-trial due to differences in movement execution mode. The second hypothesis stated that retention benefits were the result of increased movement variability produced by t-trial execution during training. The present absolute error differences found between Groups DP and DT and between Groups LP and LT support the execution mode hypothesis. Although p- and t-trial variability was equated during training through yoking, retention differences at 24 hours after training still favored the t-trial repetition groups for both distance and end-location cue recall. Thus, the variability hypothesis is not supported.

How does movement mode influence retention? As suggested earlier (Hagman, 1980b), in multitrial training situations where either p- or t-trials are emphasized through repetition, learners base later recall attempts on their retention of movement cues generated at repeated trials. It is easier to remember t-trial cues than p-trial cues because the former are learner-defined. Better retention of learner-defined cues comes from the learner's ability to predict or anticipate movement cues prior to initiation. According to Kelso (1977), "when a person is able to predict movement, two sets of signals are generated; (a) the downward discharge to effector organs, and (b) a simultaneous central discharge from motor to sensory centers that presets sensory systems for the anticipated consequences of the motor act" (p.35). Thus, the role of anticipation or prediction is to enhance the encoding of movement kinesthetic information arising from muscles and joints (Kelso, 1977; Stelmach, et. al., 1976). An extension of this corollary discharge theory can explain the superior retention resulting from t-trial repetition. It is argued that at t-trials cortical sensory centers are more prepared to receive incoming afferent impulses from muscles and joints, since movement consequences can be anticipated. At p-trials, on the other hand, this would be more difficult since little if any prior information is available regarding the terminal locus of the movement. It is this superior encoding of t-trial cues relative to p-trial cues that causes superior long-term retention.

Finally, it should be mentioned that although the present results rule out variability *per se* as the cause of t-trial retention effects, they do not rule out the possibility that variability contributes to retention, but does so only when generated during learner-defined movements. It could be argued, for example, that the effects of variability are dependent on movement mode, and perhaps vice versa. Although the present experiment does not discount this interpretation, no data have been reported either to suggest or support it. Therefore, it remains highly speculative, yet worthy of future research.

## Conclusions

The results of this experiment help to clarify past research findings and answer the question of why testing during training enhances motor

task retention. In doing so, they assist the Army in its quest to identify training methods that produce the highest levels of motor task acquisition and retention.

From the results it can be concluded that: (a) Training methods that provide for increased opportunities for testing improve long-term motor task retention; (b) these benefits derive from the superior encoding of learner-defined movements performed during t-trials, relative to experimenter-defined movements performed during p-trials; (c) increased variability of movement caused by t-trial repetition during training is not responsible for the obtained retention benefits associated with testing; (d) testing during training benefits both movement distance and end-location cue retention.

## References


Adams, J. A., & Dijkstra, S. Short-term memory for motor responses. Journal of Experimental Psychology, 1966, 71, 314-318.

Bjork, R. A. Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), Information processing and cognition the Loyola symposium. Hillsdale, N. J.: Erlbaum, 1975.

Carmer, S. G., & Swanson, M. R. An evaluation of ten pair-wise multiple comparison procedures by Monte Carlo methods. Journal of the American Statistical Association, 1973, 68, 66-74.

Guthrie, J. R. Supporting the Army in the eighties. Army Logistician, September-October 1979, 11(5), 22-25.

Hagman, J. D. Effects of Presentation- and Test-Trial Training on Motor Acquisiton and Retention. (Technical Report 431) Alexandria, Va.: US Army Research Institute for the Behavioral and Social Sciences, January, 1980(a).

Hagman, J. D. Effects of Presentation- and Test-Trial Training on Acquisition and Retention of Movement End-Location. (Technical Report 492) Alexandria, Va.: US Army Research Institute for the Behavioral and Social Sciences, November, 1980(b).

Hogan, R. M. & Kintsch, W. Differential effects of study and test trials on long-term recognition and recall. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 562-567.

Kelso, J.A.S. Planning and efferent components in the coding of movement. Journal of Motor Behavior, 1977, 9, 33-47.

Lachman, R., & Laughery, K. R. Is a test trial a training trial in free recall learning? Journal of Experimental Psychology, 1968, 76, 40-50.

Newell, K. M., & Shapiro, D. C. Variability of practice and transfer of training: Some evidence toward a schema view of motor learning. Journal of Motor Behavior, 1976, 8, 233-243.

Posner, M. I., & Keele, S. W. Retention of abstract ideas. Journal of Experimental Psychology, 1970, 83, 304-308.

Stelmach, G. E., Kelso, J. A. S., & McCullagh, P. D. Preselection and response biasing in short-term motor memory, Memory and Cognition, 1976, 4, 62-66.

Tulving, E. The effects of presentation and recall of material in free recall learning. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 175-184.

Wenger, S. K., Thompson, C. P., & Bartling, C. A. Recall facilitates subsequent recognition. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 135-144.

Wrisberg, C. A., & Schmidt, R. A. A note on motor learning without post-response knowledge of results. Journal of Motor Behavior, 1975, 7, 221-225.